UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/799,645	03/15/2004	Masayuki Yamada	00862.023518.	8043

5514 7590 09/30/2008

FITZPATRICK CELLA HARPER & SCINTO
30 ROCKEFELLER PLAZA
NEW YORK, NY 10112

EXAMINER	
COLUCCI, MICHAEL C	
ART UNIT	PAPER NUMBER
2626	
MAIL DATE	DELIVERY MODE
09/30/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/799,645	YAMADA ET AL.	
	Examiner	Art Unit	
	MICHAEL C. COLUCCI	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1,3 and 6-15* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1, 3, and 6-15* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *15 March 2004* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☒ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

NOTE: Examiner acknowledges cancellation of claims 2, 4, and 5.

Response to Arguments

1. Applicant's arguments filed 07/10/2008 have been fully considered but they are not persuasive.

Argument 1 (page 11 paragraph 3):

- "However, Shizuka and King are not seen to disclose or suggest at least the features of (i) determining whether or not an operation performed on an apparatus in a help mode designates an execution of motion, (ii) phonetically outputting a description of the motion corresponding to the operation in a case where the operation in the help mode does not designate the execution of motion, and (iii) executing the motion corresponding to the operation based on stored information, in a case where the operation in the help mode designates the execution of motion"

Response to argument 1:

Examiner takes the position that Shizuka in fact teaches the determination of whether an operation was performed, wherein a state change occurs such as being in a specific mode of operation. Shizuka teaches a system that outputs voice based on voice types which has the option of displaying a guidance mode, wherein a help mode is selected. By putting a system into a guidance mode

versus no guidance mode, a state is changed and is indicated to the system in order to properly address the new state.

Shizuka teaches a voice setting window 371 includes a drop-down list box 381 for setting the type of voice, a setting lever 382 for setting the reading speed, a setting lever 383 for setting the voice pitch for reading, a setting lever 384 for setting the strength of stress for reading, a test button 385 for reproducing a sample voice in the current voice, an OK button 386 for registering the contents that have been set and exiting the voice setting window 371, a cancel button 387 for cancelling contents that have been set and exiting the voice setting window 371, and a help button 388 for displaying, for example, a help window showing guidance of operations ([0231]).

Further, Shizuka teaches speech input mode, when help mode is not used, wherein when the camera unit 206 is rotated substantially 180 degrees by a user, in the display unit 202, a speaker 208 provided at a central portion of the back side of the camera unit 206 comes in front, as shown in FIG. 12, whereby the camera-equipped digital cellular phone 5 is switched to normal speech communication mode ([0162]).

Furthermore, Shizuka teaches a text parsing unit 306 that receives input of the text data acquired from the reading control unit 301, parses the text data to divide

it into words, and generates a phonetic symbol sequence (prosody information) with reference to dictionary data registered in the dictionary database 305 and the conversion rule registered in the conversion rule database 307, outputting it to the speech synthesis unit 308. In step S45, the speech synthesis unit 308 generates synthetic speech data based on phoneme data registered in the phoneme database 309 according to the phonetic symbol sequence supplied from the text parsing unit 306, outputting it to the speech setting unit 310. The speech setting unit 310 adjusts the synthetic speech data in accordance with the detailed speech settings that have been made using the setting levers 382 to 394 described with reference to FIG. 23, thereby generating speech data to be reproduced ([0305] – [0306]).

NOTE:

- Figure 6 of the present invention reads upon Figures 23 and 24 of Shizuka.
- Figures 1 and 2 read upon Figures 50-52 of Shizuka.

Although, Shizuka teaches state change detection with respect to phonetic speech output and operation help mode detection, the reference of King was introduced to further strengthen the prior art of Shizuka as to incorporate Shizuka in view of King. King teaches an assistive technology application 212 that produces

speech information corresponding to the screen image information. In the embodiment of FIG. 2, the speech information conveys human speech which verbally describes general attributes (e.g., color, shape, size, and the like) of the screen image and any objects (e.g., menus, dialog boxes, icons, text, and the like) within the screen image, and also includes semantic information conveying the meaning, significance, or intended purpose of each of the objects within the screen image. The speech information may include, for example, text-to-speech (TTS) commands and/or audio output signals. Suitable assistive technology applications are known and commercially available.) The assistive technology application 212 provides the speech information to a speech application program interface (API) 214. The speech application program interface (API) 214 provides a standard means of accessing routines and services within an operating system of the server 102 (King Col. 5 lines 45-65 & Fig. 2).

Further, King teaches that the console access application 202 of the client 104A are configured to cooperate such that the user of the client 104A is able to interact with the server 102 as if the user were operating the server 102 locally. As shown in FIG. 2, the client 104A includes an input device 220. The input device 220 may be for example, a keyboard, a mouse, or a voice recognition system. When the user of the client 104A activates the input device 220 (e.g., presses a keyboard key, moves a mouse, or activates a mouse button), the input

device 220 produces one or more input signals (i.e., "input signals"), and provides the input signals to the distributed console access application 202. The distributed console access application 202 transmits the input signals to the distributed console access application 200 of the server 102. (King Col. 6 lines 41-56).

Furthermore, King teaches that when the user of the client 104A is visually impaired, the user may not be able to see the screen image displayed on the display screen 210 of the client 104A. However, when the audio output device 230 produces the verbal description of the screen image, the visually-impaired user may hear the description, and understand not only the general appearance of the screen image and any objects within the screen image (e.g., color, shape, size, and the like), but also the meaning, significance, or intended purpose of any objects within the screen image as well (e.g., menus, dialog boxes, icons, and the like). This ability for a visually-impaired user to hear the verbal description of the screen image and to know the meaning, significance, or intended purpose of any objects within the screen image allows the user of the client 104A to interact with the objects in a proper, meaningful, and expected way. (King Col. 7 lines 49-64).

Shizuka in view of King together allow for a system that can detect state changes and output the changes verbally to a user, wherein a description is transmitted to

notify a user what is happening in a speech synthesis system through a verbal description that is clear enough where someone who is visually impaired can function as effectively as someone without impairment, where an audio help mode state is indicated to be turned off or on depending on the mode selected by a user where he/she will know if they are in help mode or not (King Col. 7 lines 49-64).

Claim Rejections - 35 USC § 103

2. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

3. Claim 1, 3, 6-8 and 13-15 are rejected under 35 U.S.C. 103(a) as being unpatentable over Shizuka et al. US 20020184004 A1 (hereinafter Shizuka) King et al. US 7103551 B2 (hereinafter King).

Re claims 1 and 13-15, Shizuka teaches a data processing method comprising:

an operation detection step of detecting operation performed on an apparatus ([0395]);

a state determination step of determining whether a state of the apparatus is a normal mode or a help mode when the operation is detected in said operation detection step ([0235]);

a first execution step of executing motion corresponding to the operation in a case where it is determined in said state determination step that the state of the apparatus is the normal mode ([0231]);

an execution determination step of determining whether or not the operation detected in said operation detection step designates an execution of motion in a case where it is determined in said state determination step that the state of the apparatus is the help mode ([0235]);

an audio output step of phonetically outputting ([0305] – [0306]) a description of the motion corresponding to the operation in a case where it is determined in said execution determination step that the operation detected in said operation detection step does not designate the execution of motion ([0231]);

a storage step of storing in a storage device information regarding the operation ([0234] – [0235]),

a second execution step of executing the motion ([0231]) corresponding to the operation based on the information stored in the storage device ([0239]), in a case where it is determined in said execution determination step that the operation detected in said operation detection step designates the execution of motion ([0240]).

However, Shizuka fails to teach outputting a description of the motion corresponding to the operation in a case where the state of the apparatus is the help mode;

King teaches an assistive technology application 212 that produces speech information corresponding to the screen image information. In the embodiment of FIG.

2, the speech information conveys human speech which verbally describes general attributes (e.g., color, shape, size, and the like) of the screen image and any objects (e.g., menus, dialog boxes, icons, text, and the like) within the screen image, and also includes semantic information conveying the meaning, significance, or intended purpose of each of the objects within the screen image. The speech information may include, for example, text-to-speech (TTS) commands and/or audio output signals. Suitable assistive technology applications are known and commercially available.) The assistive technology application 212 provides the speech information to a speech application program interface (API) 214. The speech application program interface (API) 214 provides a standard means of accessing routines and services within an operating system of the server 102 (King Col. 5 lines 45-65 & Fig. 2).

Further, King teaches that the console access application 202 of the client 104A are configured to cooperate such that the user of the client 104A is able to interact with the server 102 as if the user were operating the server 102 locally. As shown in FIG. 2, the client 104A includes an input device 220. The input device 220 may be for example, a keyboard, a mouse, or a voice recognition system. When the user of the client 104A activates the input device 220 (e.g., presses a keyboard key, moves a mouse, or activates a mouse button), the input device 220 produces one or more input signals (i.e., "input signals"), and provides the input signals to the distributed console access application 202. The distributed console access application 202 transmits the input signals to the distributed console access application 200 of the server 102. (King Col. 6 lines 41-56).

Furthermore, King teaches that when the user of the client 104A is visually impaired, the user may not be able to see the screen image displayed on the display screen 210 of the client 104A. However, when the audio output device 230 produces the verbal description of the screen image, the visually-impaired user may hear the description, and understand not only the general appearance of the screen image and any objects within the screen image (e.g., color, shape, size, and the like), but also the meaning, significance, or intended purpose of any objects within the screen image as well (e.g., menus, dialog boxes, icons, and the like). This ability for a visually-impaired user to hear the verbal description of the screen image and to know the meaning, significance, or intended purpose of any objects within the screen image allows the user of the client 104A to interact with the objects in a proper, meaningful, and expected way. (King Col. 7 lines 49-64).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Shizuka to incorporate outputting a description of the motion corresponding to the operation in a case where the state of the apparatus is the help mode and a description of the motion corresponding to the operation in a case where it is determined in said execution determination step that the operation detected in said operation detection step does not designate the execution of motion as taught by King to allow for a system that can detect state changes and output the changes verbally to a user, wherein a description is transmitted to notify a user what is happening in a speech synthesis system through a verbal description that is clear enough where someone who is visually impaired can function as effectively as someone

without impairment, where an audio help mode state is indicated to be turned off or on depending on the mode selected by a user where he/she will know if they are in help mode or not (King Col. 7 lines 49-64).

Re claim 3, Shizuka teaches the data processing method according to claim 1, further comprising:

a cancellation step of canceling the help mode of the apparatus in a case where the state of the apparatus is the help mode and said operation is help operation ([0231]);

a setting step of setting the state of the apparatus in the help mode in a case where the state of the apparatus is not the help mode and said operation is help operation ([0231] & Fig. 23).

Re claim 6, Shizuka teaches the data processing method according to claim 1, further comprising a termination step of terminating audio output being currently outputted in a case where operation performed on the apparatus is detected in said operation detection step (Fig. 34 items S48 and S49).

Re claim 7, Shizuka teaches the data processing method according to claim 1, further comprising a second audio output step of phonetically outputting a motion result of said operation executed in said second execution step ([0305] – [0306]).

Re claim 8, Shizuka teaches the data processing method according to claim 1, further comprising:

an acquisition step of acquiring a name of said operation performed on the apparatus ([0225]);

However, a third audio output step of phonetically outputting the name before phonetically outputting the description of the motion in said audio output step (King Col. 7 lines 49-64);

King teaches an assistive technology application 212 that produces speech information corresponding to the screen image information. In the embodiment of FIG. 2, the speech information conveys human speech which verbally describes general attributes (e.g., color, shape, size, and the like) of the screen image and any objects (e.g., menus, dialog boxes, icons, text, and the like) within the screen image, and also includes semantic information conveying the meaning, significance, or intended purpose of each of the objects within the screen image. The speech information may include, for example, text-to-speech (TTS) commands and/or audio output signals. Suitable assistive technology applications are known and commercially available.) The assistive technology application 212 provides the speech information to a speech application program interface (API) 214. The speech application program interface (API) 214 provides a standard means of accessing routines and services within an operating system of the server 102 (King Col. 5 lines 45-65 & Fig. 2).

Further, King teaches that the console access application 202 of the client 104A are configured to cooperate such that the user of the client 104A is able to interact with

the server 102 as if the user were operating the server 102 locally. As shown in FIG. 2, the client 104A includes an input device 220. The input device 220 may be for example, a keyboard, a mouse, or a voice recognition system. When the user of the client 104A activates the input device 220 (e.g., presses a keyboard key, moves a mouse, or activates a mouse button), the input device 220 produces one or more input signals (i.e., "input signals"), and provides the input signals to the distributed console access application 202. The distributed console access application 202 transmits the input signals to the distributed console access application 200 of the server 102. (King Col. 6 lines 41-56).

Furthermore, King teaches that when the user of the client 104A is visually impaired, the user may not be able to see the screen image displayed on the display screen 210 of the client 104A. However, when the audio output device 230 produces the verbal description of the screen image, the visually-impaired user may hear the description, and understand not only the general appearance of the screen image and any objects within the screen image (e.g., color, shape, size, and the like), but also the meaning, significance, or intended purpose of any objects within the screen image as well (e.g., menus, dialog boxes, icons, and the like). This ability for a visually-impaired user to hear the verbal description of the screen image and to know the meaning, significance, or intended purpose of any objects within the screen image allows the user of the client 104A to interact with the objects in a proper, meaningful, and expected way. (King Col. 7 lines 49-64).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Shizuka to incorporate outputting a description of the motion corresponding to the operation in a case where the state of the apparatus is the help mode and a description of the motion corresponding to the operation in a case where it is determined in said execution determination step that the operation detected in said operation detection step does not designate the execution of motion as taught by King to allow for a system that can detect state changes and output the changes verbally to a user, wherein a description is transmitted to notify a user what is happening in a speech synthesis system through a verbal description that is clear enough where someone who is visually impaired can function as effectively as someone without impairment, where an audio help mode state is indicated to be turned off or on depending on the mode selected by a user where he/she will know if they are in help mode or not (King Col. 7 lines 49-64).

4. Claims 9-12 are rejected under 35 U.S.C. 103(a) as being unpatentable over Shizuka et al. US 20020184004 A1 (hereinafter Shizuka) King et al. US 7103551 B2 (hereinafter King) and further in view of Surace et al. US 6334103 B1 (hereinafter Surace).

Re claim 9, Shizuka teaches the data processing method according to claim 1, further comprising:

a changing step of changing sound quality of output speech (Fig. 24)

However, Shizuka in view of King fails to teach a determination step of determining whether or not one same operation has been repeatedly performed on the apparatus (Surace Col. 10 lines 22-30);

from the speech outputted last, in a case where one same operation has been repeatedly performed (Surace Col. 10 lines 22-30).

Surace teaches a voice user interface with personality, wherein it is determined whether the user is requiring repeated help in the same session or across sessions (i.e., a user is requiring help more than once in the current session).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Shizuka in view of King to incorporate changing sound quality after determining whether or not the same operation has been repeatedly performed as taught by Surace to allow for a voice quality adjustment to be implemented such as a personality of a user interface dependent on how many times an operation is repeated (based on social and psychological experimental data) (Surace Col. 10 lines 22-30).

Re claim 10, data processing method according to claim 9, wherein in said changing step, vocalize speed of the output speech is changed (Fig. 24).

Re claim 11, Shizuka in view of King fails to teach the data processing method according to claim 9, wherein in said changing step, volume of the output speech is changed (Surace Col. 22 lines 44-49).

Surace teaches the editing of audio tapes of the recorded scripts (e.g., to adjust volume and ensure smooth audio transitions within dialogs).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Shizuka in view of King to incorporate changing the volume of the output speech as taught by Surace to allow for a voice quality adjustment to be implemented such as a personality of a user interface dependent on how many times an operation is repeated (based on social and psychological experimental data). Various parameters such as pitch, speed, clarity, and intonation can be varied to alter the personality of a voice interface (Surace Col. 22 lines 44-49).

Re claim 12, Shizuka teaches the data processing method according to claim 9, wherein in said changing step, vocal quality of the output speech is changed (Fig. 24).

Conclusion

5. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any

extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Michael C. Colucci whose telephone number is (571)-270-1847. The examiner can normally be reached on 9:30 am - 6:00 pm, Monday-Friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (571)-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Michael C Colucci/
Examiner, Art Unit 2626
Patent Examiner
AU 2626
(571)-270-1847
Michael.Colucci@uspto.gov

/Richemond Dorvil/
Supervisory Patent Examiner, Art Unit 2626